RESEARCH ARTICLE

Funding COVID-19 research: Insights from an exploratory analysis using open data infrastructures

Alexis-Michel Mugabushaka[1], Nees Jan van Eck[2], and Ludo Waltman[2]

[1]European Research Council, Brussels, Belgium

[2]Centre for Science and Technology Studies (CWTS), Leiden University, Leiden, Netherlands

Keywords: COVID-19, Crossref, funding data, open metadata, Scopus, Web of Science

ABSTRACT

To analyze the outcomes of the funding they provide, it is essential for funding agencies to be able to trace the publications resulting from their funding. We study the open availability of funding data in Crossref, focusing on funding data for publications that report research related to COVID-19. We also present a comparison with the funding data available in two proprietary bibliometric databases: Scopus and Web of Science. Our analysis reveals limited coverage of funding data in Crossref. It also shows problems related to the quality of funding data, especially in Scopus. We offer recommendations for improving the open availability of funding data in Crossref.

an open access journal

Citation: Mugabushaka, A.-M., van Eck, N. J., & Waltman, L. (2022). Funding COVID-19 research: Insights from an exploratory analysis using open data infrastructures. *Quantitative Science Studies*, *3*(3), 560–582. https://doi.org/10.1162/qss_a_00212

DOI: https://doi.org/10.1162/qss_a_00212

Peer Review: https://publons.com/publon/10.1162/qss_a_00212

Received: 7 March 2022
Accepted: 29 July 2022

Corresponding Author:
Alexis-Michel Mugabushaka
Alexis-Michel.MUGABUSHAKA@ec.europa.eu

Handling Editor:
Vincent Larivière

1. INTRODUCTION

The ongoing coronavirus 2019 pandemic (COVID-19) has caused a major public crisis. It has advanced to a major cause of death and overwhelmed healthcare systems in many countries. According to the World Health Organization (WHO) COVID-19 dashboard, as of January 2022, there have been over 340 million confirmed cases and 5.5 million deaths worldwide. The measures taken to contain its spread have also caused unprecedented disruptions of economic and social life around the world.

Researchers have been among the "first responding" professions dealing with the pandemic and its consequences. They advise public authorities on the best measures to control the pandemic, study the course of the disease, develop clinical guidelines and medical protocols, and very importantly develop vaccines—some of them in record time—as well as therapies.

Around the world, several research teams have redirected their research efforts to help fight the pandemic (Hao, 2020; Kwon, 2020; Viglione, 2020). Research funding bodies have multiplied initiatives to support research related to the pandemic. In addition to measures allowing grant management flexibility (Stoye, 2020), several organizations have launched fast-track research funding programs specifically targeted at various aspects of the crisis. For example, in the United States, the National Institutes of Health (NIH) launched several initiatives to tackle the pandemic by using existing funding mechanisms or establishing new, dedicated programs. They are bundled in the NIH-Wide Strategic Plan for COVID-19 Research. The National Science Foundation (NSF) activated its Rapid Response Research mechanisms (RAPID) used for research funding in unanticipated events. In Europe, the European Union

The MIT Press

launched a COVID-19 emergency call for proposals in January 2020 and published subsequent calls throughout the year. The Innovative Medicines Initiative (IMI) launched a fast-track call for proposals to speed up the development of new drugs and diagnostics to halt the global outbreak of COVID-19. The section "national activities" of the European Research Area (ERA) corona platform lists several other initiatives launched at the beginning of the pandemic[1]. They include the Deutsche Forschungsgemeinschaft (DFG—German Research Foundation), which set up a COVID-19-focused funding program, and the Swedish Research Council, which, among other initiatives, teamed up with the National Natural Science Foundation of China (NSFC) to support collaborative projects on coronavirus.

The Organisation for Economic Co-operation and Development's COVID-19 WATCH, which monitors research policy responses to the COVID-19 crisis, estimates the combined value of public research funding in those measures to be about US$2.6 billion, and US$3.8 billion if other sources (charities, industry) are also considered.

This has led to a massive expansion of COVID-19 research. Scientific publishers have adapted their editorial processes to allow fast dissemination of new results (Hurst & Greaves, 2021), posing to researchers and the public a particular "challenge of discerning signal amidst noises," as the editors of one journal put it (Bleck, Buchman et al., 2020).

The resulting unprecedented increase of research papers on a single topic—by some accounts over 100,000 in 2020 alone, accounting for about 4% of total research outputs (Else, 2020)—has also triggered a large body of metaresearch on COVID-19 research. One strand of this research seeks to tame what has been termed a "paper tsunami" (Brainard, 2020). It uses advanced machine learning techniques for information extraction, misinformation detection, question answering, etc. (Shorten, Khoshgoftaar, & Furht, 2021). Another line of work uses scientometric techniques to develop an understanding of the output of COVID-19 publications. This line of research, for instance, studies the role of countries, institutions, journals, and authors (Mohadab, Bouikhalene, & Safi, 2020; Tao, Zhou et al., 2020), specific fields and techniques (Aristovnik, Ravšelj, & Umek, 2020; Hossain, Sarwar et al., 2020), gender (Andersen, Nielsen et al., 2020), research areas (Colavizza, Costas et al., 2021) and researchers (Ioannidis, Salholz-Hillel et al., 2021).

One aspect that remains underexplored is how this research has been, and is being, funded. One of the notable exceptions is the recent work by Cross, Rho et al. (2021) entitled "Who funded the research behind the Oxford-AstraZeneca COVID-19 vaccine?" In this work, the authors analyzed funding information of about 100 peer-reviewed articles relevant to the Chimpanzee adenovirus-vectored vaccine (ChAdOx) on which the Oxford-AstraZeneca vaccine is based. The authors found that this research was almost entirely supported by public funding. The European Commission, the Wellcome Trust, and the Coalition for Epidemic Preparedness Innovations (CEPI) were the biggest funders of ChAdOx research and development. The study also highlights the lack of transparency in reporting of funding, which "hinders the discourse surrounding public and private contributions towards R&D and the cost of R&D."

Another study in this context is the analysis of how NIH funding has contributed to research used in COVID-19 vaccine development (Kiszewski, Cleary et al., 2021). The authors focused on 10 technologies employed in candidate vaccines (as of July 2020), identified from WHO documents, and on research on five viruses with epidemic potential. They then estimated the NIH funding to those areas by linking relevant publications (identified by searching in PubMed via MeSH terms) to grants using acknowledgments. The authors concluded that NIH funding has significantly contributed to advances in vaccine technologies, which helped the rapid

[1] https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/covid-19?tabId=5.

development of COVID-19 vaccines. However, they also noted that NIH funding for vaccine research for specific pandemic threats has been inconsistent and called for sustained public sector funding for better preparedness against future pandemics.

In this paper, we expand the scope from a single technology or single funding body to research on COVID-19 in general. Getting an accurate picture of COVID-19 research funding is important for a number of reasons:

- Insights into the various funding mechanisms and modalities used and how (relatively) successful they have been may inform the organization of funding in case of future emergencies.
- Given the societal interest and policy implications of the outcomes of COVID-19 research, it is important to understand how these outcomes relate to the interests of sponsors. Most publication ethics guidelines require researchers to state the role of funders in reported research. Although this is mostly applied in medical journals, extending it to research on other aspects of the pandemic can bring the transparency needed to assess the credibility of scientific findings.
- Most, if not all, public research organizations funding COVID-19 research have accountability obligations. They must report to public authorities on the results of their funding activities. Studying the funding patterns of this research can help funders understand not only the results of their activities but also how these results relate to research funded by others and, by putting it in a wider perspective, their relative weight in funding COVID-19 research.
- The concerns over fair vaccine access and the resulting debates on patent waivers for COVID-19 vaccines could be better informed by reliable data on the public investments that enabled the vaccine development.

In this paper, we explore the funding of COVID-19 research. The main objective is to find out which funding organizations have contributed to COVID-19 research reported in the scholarly literature. We seek specifically

- to explore the extent to which funding data can be found in openly available databases, in particular Crossref;
- to identify the main funding organizations that supported COVID-19 research; and
- to compare the findings based on openly available databases with those based on proprietary databases.

Another study of COVID-19 research funding was carried out by Shueb, Gul et al. (2022). Unlike our work reported in the present paper, the study by Shueb et al. covers only research published in the first months of the pandemic and does not make use of openly available funding data.

This paper is organized as follows. In Section 2, we discuss the data used in our analyses. We report our results based on openly available data in Section 3 and present a comparison with results based on proprietary data in Section 4. We summarize our findings and draw conclusions in Section 5.

2. DATA

We use the funding acknowledged by publications as a proxy for funding of the underlying research (Álvarez-Bornstein & Montesi, 2020). This requires combining data on COVID-19 related publications and data on the funding sources of these publications. In this section, we discuss the data we combined to link publications to funding as well as the data resulting from this linking.

2.1. CORD-19 Data

We use the COVID-19 Open Research Dataset (CORD-19), a data set of COVID-19 research articles (both metadata and full text) released by Semantic Scholar in partnership with other organizations. CORD-19 defines itself as "a comprehensive collection of publications and preprints on COVID-19 and related historical coronaviruses such as SARS and MERS" (Wang, Lo et al., 2020).

CORD-19 combines data from different sources which follow mainly the same search approach. It is updated regularly by adding new records and deleting erroneous or retracted entries. While CORD-19 is the most widely used COVID-19 literature data set, it is not without limitations. The search approach used by CORD-19 has the advantage of conceptual clarity (i.e., papers included say something about the last three outbreaks caused by coronaviruses or about coronaviruses more generally), but this advantage is also its inherent limitation: keywords-based search may lead to the inclusion of papers which only cursorily mention a coronavirus outbreak (false positive) or miss relevant papers which use other terms (false negatives). Other limitations, acknowledged by the CORD-19 team, include the restriction of the data to scholarly publications, including preprints, leaving aside "*other types of documents that could be important, such as technical reports, white papers, informational publications by governmental bodies*" (Wang et al., 2020) as well as the focus on English language publications.

Some research has critically inspected the CORD-19 data set with respect to coverage and has suggested possible improvements. Kanakia, Wang et al. (2020) explored how citation links can be used to understand and mitigate possible bias in CORD-19. Colavizza et al. (2021) also studied the coverage of CORD-19, using a version of the data set from July 2020. Within the data set, they identified a subset called "CORD-19 strict," for which the CORD-19 query matches the title and abstract of a publication, disregarding the full text. They found that this subset of CORD-19 almost perfectly matches the results retrieved from the Web of Science (WoS) database, indicating that CORD-19 "provides an almost complete coverage of research on COVID-19 and coronaviruses." However, the fact that this subset is small suggests that CORD-19 does not only cover COVID-19 research, leading Colavizza et al. to caution users to be aware that CORD-19 may include "a large number of publications whose relevance for COVID-19 and coronaviruses research needs a more careful assessment, and some of which may be of limited relevance." The uncertainties in the scope of CORD-19 and inevitable errors due to its data collection approach are a limitation of our analysis that should be kept in mind when interpreting our results.

In the rest of this paper, we refer to publications in the CORD-19 data set as *COVID-19 publications*. This should be understood as publications that are in a broad sense related to COVID-19, including publications that appeared before the start of the COVID-19 pandemic and that deal with coronavirus research more broadly.

We used the CORD-19 data set released by the Allen Institute on 15 February 2021, in the version enriched by Microsoft Academic (MAG) by adding publication identifiers from MAG.

2.2. Crossref Data

Linking publications to funding sources is far from straightforward. As discussed elsewhere (Mugabushaka, 2020), several approaches can be used, each with their advantages and limitations. Our primary focus is on funding data provided by Crossref, although we also perform a comparison with funding data from proprietary databases.

2.2.1. Crossref funding data

Crossref is a not-for-profit organization that provides an open infrastructure used by many stakeholders in the scholarly communication system. Its members include publishers, universities, preprint services, and funding organizations. Its primary function is to enable parties globally to update and exchange metadata about the scholarly record, identified through Digital Object Identifiers (DOIs) and made open for all.

Crossref encourages its members to deposit metadata going beyond standard bibliographic information. Those rich metadata may also include funding data. According to the guidance Crossref gives to its members, funding data can be obtained from authors when they submit a manuscript or extracted from the acknowledgment or funding information section of a manuscript. As part of its data curation, Crossref can also add missing data, for example by inferring missing funder identifiers from funders names. The share of Crossref records with funding data has steadily increased to reach about 25% in 2019 (Hendricks, Tkaczyk et al., 2020, Figure 3).

Crossref makes data on funding together with other publication metadata openly available. We use Crossref's XML Metadata Plus Snapshot. The snapshot was downloaded on March 5, 2021. We consider only the 110,851,607 records classified as journal article, book content, conference paper, or preprint[2].

2.2.2. Crossref Funder Registry

Funding data in Crossref are powered by a taxonomy of funders maintained in the Crossref Funder Registry. The Funder Registry was started by Elsevier in 2012 and was donated to Crossref. The curation of the registry is supported by Elsevier, which reviews it every 4–6 weeks to add new funding entities as well as update or correct existing ones.

The Funder Registry assigns a unique identifier, a DOI, to each funder. The registry is organized in a hierarchy in which individual entries are linked to parent and child entries.

In Crossref, funding data may refer to any hierarchical level in the Funder Registry. To give two examples

- In the case of the NIH, funding data may refer to a specific institute such as the National Institute of Allergy and Infectious Diseases (DOI: 10.13039/100000060), the NIH as a whole (DOI: 10.13039/100000002), or the US Department of Health and Human Services (DOI: 10.13039/100000016).
- In the case of the European Union, funding data may refer to the Marie Skłodowska-Curie program (DOI: 10.13039/100010665), the European Research Council (DOI: 10.13039/100011199), the H2020 program (DOI: 10.13039/501100007601), or the European Commission (DOI: 10.13039/501100000780).

Funding data captured by publishers and submitted to Crossref reflects the different acknowledgment practices of authors. This can lead to an inconsistent picture of the contributions of funders of COVID-19 research. For a more accurate picture, there is a need to group funders based on the hierarchy of funding organizations in the Funder Registry. Because this

[2] In late 2019, funding organizations started registering grant metadata in Crossref, which is then linked to publications. In this paper, we focus on funding data submitted to Crossref by publishers.

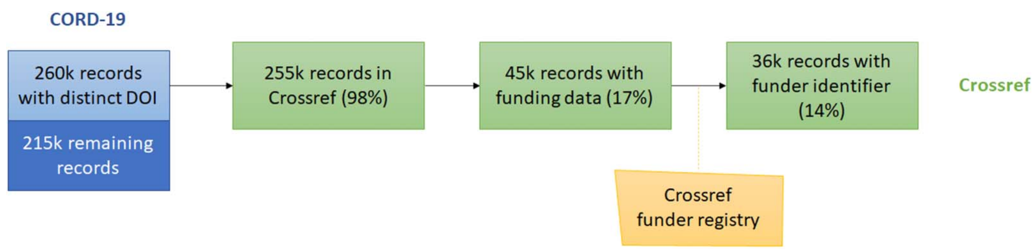

Figure 1. Funding data for COVID-19 publications in Crossref.

hierarchy follows the legal structure of funding bodies, it can, in some cases, make comparisons difficult. For example, public funding bodies in Canada have the Government of Canada at the highest level in the hierarchy of the Funder Registry, and in the United States government departments usually constitute the highest level. We created a mapping of each entry in the Funder Registry to the corresponding top-level entity in the Funder Registry hierarchy (van Eck & Mugabushaka, 2021). The mapping is based on version 1.34 of the Funder Registry, which has 27,741 entries at lower levels, grouped into 22,369 funders at the highest level. While not perfect, this approach has the advantage of transparency and simplicity. Creating an alternative mapping would not only require detailed knowledge of funding structures worldwide but also require subjective choices that might bias our analysis.

2.3. Linking the Data Sets

Figure 1 illustrates how the data sets were linked. We used DOIs to link CORD-19 publications to publications in Crossref. The CORD-19 version we used includes 484,064 records, of which 474,691 are unique publication records (i.e., unique CORD-19 identifiers). Of these records, 260,636 (or about 55%) have a DOI in CORD-19. After eliminating duplicates, we ended up with 259,652 unique DOIs, of which 255,378 were found in Crossref. Our analysis is based on these 255,378 records in Crossref.

The lack of DOIs for a substantial share of the publications in CORD-19—including about half of the publications in 2020 and 2021—is another important limitation of our analysis. It means that the results reported in the subsequent sections offer only a partial picture of the funding of COVID-19 research.

Of the CORD-19 publications linked to Crossref, 44,820 have funding data. For 36,008 publications, we also have an identifier of a funding organization included in the Crossref Funder Registry. Our analysis of COVID-19 funding is based on these 36,008 publications.

The relatively low share of publications with funding data in Crossref is another limitation of our analysis. An important implication of this limitation is the need to pay attention to the way in which funding data is collected, and which measures can be taken by various stakeholders to increase the availability of funding data in open data infrastructures. We will share some reflections and suggestions in the concluding section.

3. ANALYSIS OF OPEN FUNDING DATA IN CROSSREF

The data at hand, effectively a relatively limited subset of COVID-19 research papers due to the data limitations described above, offers an incomplete picture of COVID-19 funding. As partial as the results are, however, they can give indications of funding patterns, notably for the most prolific funders and how they interact in the network of "cofunding." The results also provide insight into the type of funding bodies supporting COVID-19 research.

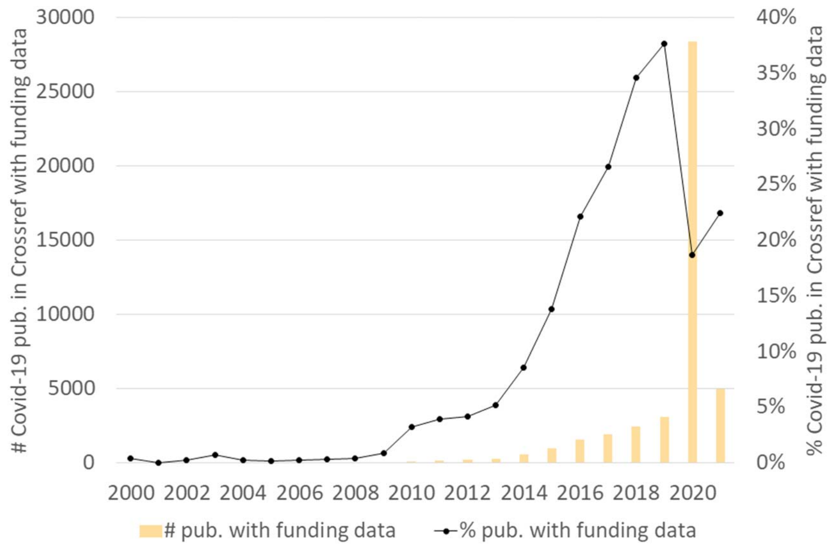

Figure 2. Percentage of COVID-19 publications with a DOI that have funding data in Crossref.

3.1. Availability of Funding Data

As noted above, we focus on publications indexed in Crossref, as our aim is to use open funding data. Funding data are available in Crossref for 44,820 COVID-19 publications. This accounts for 17% of the COVID-19 publications for which a DOI is available.

As Figure 2 shows, the availability of funding data in Crossref has increased over time, reaching almost 40% in 2019. This is in line with earlier analyses showing that the amount of funding data submitted by publishers to Crossref has steadily increased (Hendricks et al., 2020; van Eck & Waltman, 2021). Before 2020, the share of COVID-19 publications with funding data is higher than the overall share of publications with funding data. The sharp decline in 2020 in the share of publications with funding data may seem puzzling. It could be that at the beginning of the COVID-19 pandemic many researchers immediately started to work on COVID-19-related research projects, without first applying for funding. This impression is confirmed when looking at other databases. Both WoS and Scopus also show a significant decrease in the share of publications with funding data, reinforcing the idea that a relatively large share of all COVID-19 research in 2020 did not receive funding from a funding agency.

3.2. Top Funders

The 36,008 publications linked to a funder identifier acknowledge 5,386 distinct funders at the top level of the Crossref Funder Registry hierarchy. This indicates that, based on open funding data in Crossref, close to one in four funding bodies at the top level in the Funder Registry have supported COVID-19 research. The number of papers per funder varies significantly. The median is two papers. The top 10 funders account for over half of all papers.

Table 1 shows the top 30 funders with the largest number of COVID-19 publications. Based on open funding data in Crossref, the top funders are the US Department of Health and Human Services, with 7,081 publications and the National Natural Science Foundation of China with 5,318 publications. They account for 20% and 15%, respectively, of the papers for which funding data is available. They are followed by the European Commission in third place with 1,313 publications and by UK Research and Innovation and the Ministry of Science and Technology of the People's Republic of China, with 1,111 and 1,128 publications, respectively. Other funders in the top 10 are the Government of Canada, the U.S. National Science

Table 1. Top 30 funders of COVID-19 publications (based on Crossref data)

Country	Funder	Number of publications
USA	U.S. Department of Health and Human Services	7,081
CHN	National Natural Science Foundation of China	5,318
EU	European Commission	1,313
CHN	Ministry of Science and Technology of the People's Republic of China	1,128
GBR	UK Research and Innovation	1,111
CAN	Government of Canada	970
USA	National Science Foundation	942
DEU	Deutsche Forschungsgemeinschaft	730
JPN	Ministry of Education, Culture, Sports, Science and Technology	691
GBR	National Institute for Health Research	655
BRA	Ministério da Ciência, Tecnologia e Inovação	560
KOR	National Research Foundation of Korea	536
USA	US Department of Defense	531
AUS	Department of Health, Australian Government	498
USA	Bill and Melinda Gates Foundation	473
CHN	Ministry of Education of the People's Republic of China	466
BRA	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior	450
ESP	Ministerio de Economía y Competitividad	418
CHN	Ministry of Finance	343
GBR	Wellcome Trust	340
CHN	China Postdoctoral Science Foundation	298
BRA	Fundação de Amparo à Pesquisa do Estado de São Paulo	273
FRA	Agence Nationale de la Recherche	271
CHE	Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung	265
TWN	Ministry of Science and Technology, Taiwan	261
JPN	Japan Agency for Medical Research and Development	256
USA	US Department of Agriculture	251
IND	Department of Science and Technology, Ministry of Science and Technology, India	236
USA	Foundation for the National Institutes of Health	229
DEU	Bundesministerium für Bildung und Forschung	223

Foundation, the Deutsche Forschungsgemeinschaft, the Japanese Ministry of Education, Culture, Sports, Science and Technology, and the UK National Institute for Health Research.

3.3. Type of Funder

For each funding body, the Crossref Funder Registry includes information about the type of organization. This information is organized in two dimensions. On the one hand, a distinction is made between private and public organizations. On the other hand, for each of these categories, a further distinction is made between different organization forms.

Table 2 shows that over two-thirds of the publications (28,186) acknowledge funding bodies classified as "government" while the rest (13,900) acknowledge funding from private entities.

The classification of funding organizations in the Funder Registry follows the legal status of these organizations in different countries. Given their particularities, this can lead to results that are difficult to compare. For example, one of the major public funding bodies in Germany, the DFG, is classified in the Funder Registry as a private organization under "trusts, charities, foundations," which is indeed its legal form. In a way, however, it is comparable to other funders classified as "government"—such as the NIH and NSF in the United States—as it receives its funding from public authorities (at both the federal and local levels).

The classification of organizations in the Funder Registry allows us to identify other nonpublic players active in funding COVID-19 research. Among those with more than 100 publications, we find philanthropic organizations such as the Wellcome Trust and the Bill & Melinda Gates Foundation and pharmaceutical companies such as Pfizer, Sanofi, and Novartis (see "Sec_3_2" in the Supplementary material; Mugabushaka, van Eck, & Waltman, 2022).

3.4. Cofunding Rate and Cofunding Network

As shown in Figure 3, one-third of the COVID-19 papers with funding data are linked to more than one funding body. In these cases, the research team behind the reported work may have multiple lines of funding that contributed to the reported results, or the authors may belong to multiple research teams with different sources of funding.

Table 2. Number of COVID-19 publications by type of funding organization (based on Crossref data)

Type of funding organization	Public	Private	All
National government	23,959		23,959
Trusts, charities, foundations (both public and private)	10	6,355	6,365
Local government	6,309		6,309
Universities (academic only)	843	3,732	4,575
Other nonprofit organizations	18	3,073	3,091
For-profit companies (industry)	577	1,204	1,781
Associations and societies (private and public)	55	987	1,042
International organizations	6	742	748
Research institutes and centers	59	274	333
Total	31,836	16,367	48,203
Distinct total	28,186	13,900	42,086

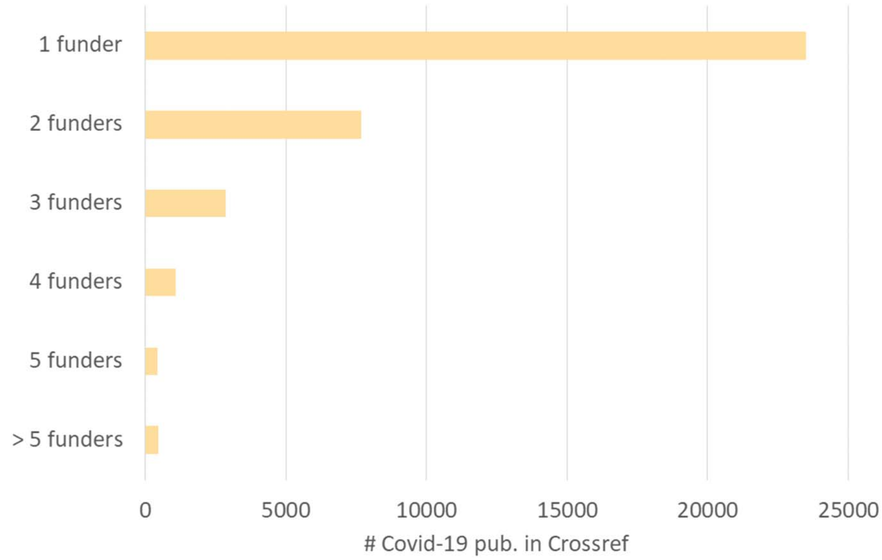

Figure 3. Number of COVID-19 publications by number of funders (based on Crossref data).

Through papers acknowledging multiple sources of funding, funding bodies are effectively engaged in a cofunding collaboration network. Looking at the funders with the largest number of publications (see "Sec_3_4b" in the Supplementary material; Mugabushaka et al., 2022), we see that the share of publications in which a funding agency is acknowledged together with other funders is, on average, relatively high, but with some variation across funders. For some of the large funders, the share of publications cofunded with other funders is around 50%. Examples are the National Natural Science Foundation of China (48%), the Government of Canada (55%), the Deutsche Forschungsgemeinschaft (55%), and the US Department of Health and Human Services (58%). For other large funders, such as UK Research and Innovation (68%) and the European Commission (72%), more than two-thirds of the publications are cofunded with other funders.

Figure 4 presents a visualization of a cofunding network for COVID-19 publications. The visualization was created using the VOSviewer software (van Eck & Waltman, 2010). The network includes 384 funders that each have at least 30 cofunding links with other funders in the network. The visualization can be explored interactively at https://tinyurl.com/z27f97ek.

4. COMPARISON WITH PROPRIETARY FUNDING DATA

To assess the comprehensiveness of open funding data in Crossref, we performed a comparison with funding data in two proprietary bibliometric databases: Scopus (Baas, Schotten et al., 2020) and WoS (Birkle, Pendlebury et al., 2020). Following the approach taken for Crossref, we use the subset of the CORD-19 data set with DOIs to query Scopus and WoS and retrieve funding data. The comparison with funding data in proprietary databases aims to provide insight into the comprehensiveness of open funding data in Crossref. Comparing the funding data made available by different proprietary databases is not the main purpose of our analysis. For earlier analyses of funding data available in Scopus and WoS, we refer to Álvarez-Bornstein, Morillo, and Bordons (2017), Grassano, Rotolo et al. (2017), Kokol and Blažun Vošner (2018), Liu (2020), Liu, Tang, and Hu (2020), Paul-Hus, Desrochers, and Costas (2016), and Tang, Hu, and Liu (2017).

To have a meaningful comparison, we focus on funding data obtained from publishers and made available in bibliometric databases. We do not consider funding data collected from

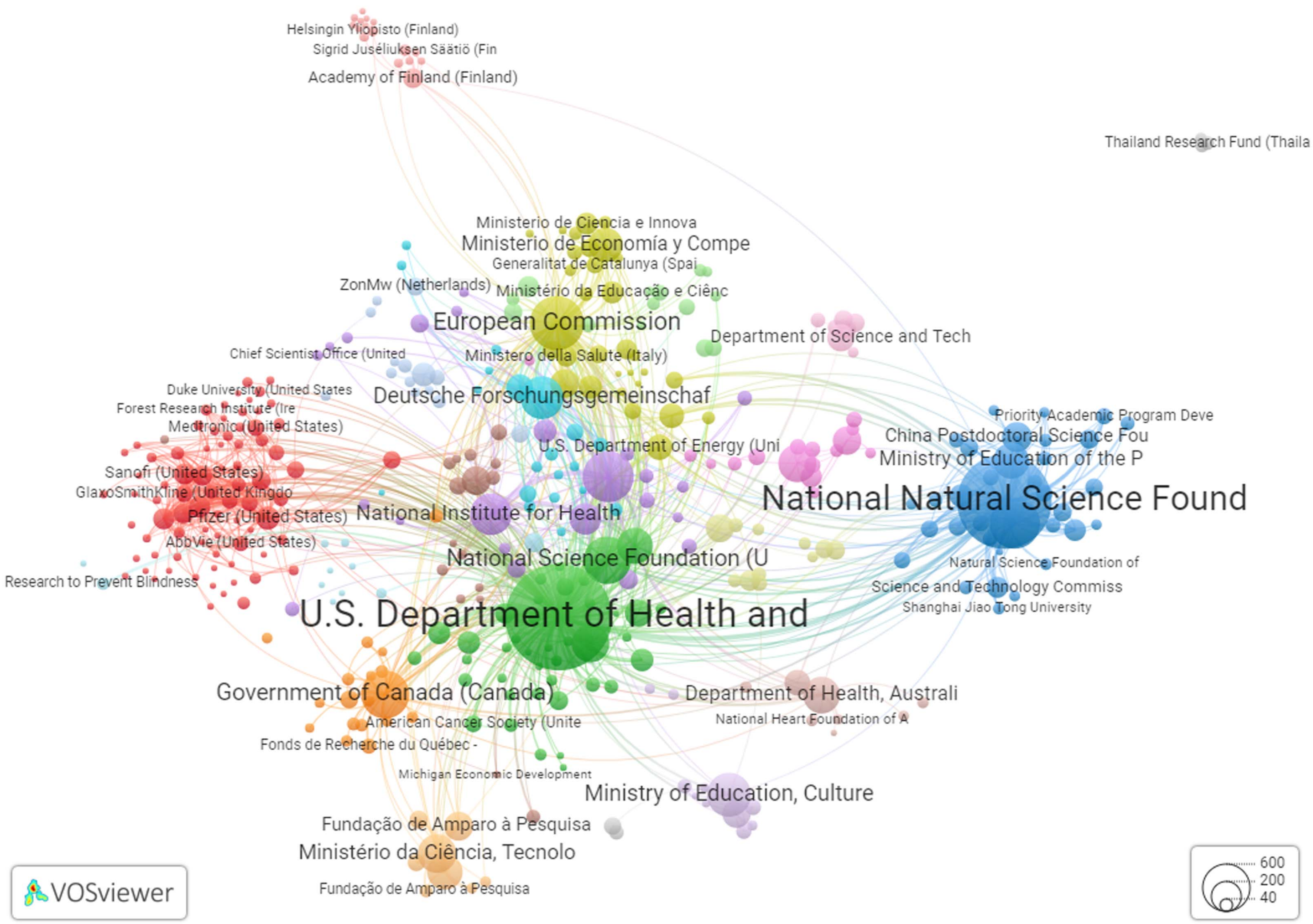

Figure 4. Co-funding network for COVID-19 publications (based on Crossref data).

funding agencies. Considering data obtained from funding agencies would obscure the comparison between Crossref and proprietary bibliometric databases because our analysis for Crossref takes into account only data obtained from publishers. WoS nowadays also includes funding data obtained from funding agencies, such as data from NIH Reporter, but we do not use this data. We also do not consider funding data from the Dimensions database (Herzog, Hook, & Konkiel, 2020), as this database does not make a distinction between funding data obtained from publishers and from funding agencies.

The Scopus and WoS data were retrieved from the in-house database system of the Centre for Science and Technology Studies (CWTS) at Leiden University. For both databases, we used data from April 2021. The following WoS citation indexes were used: Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (AHCI), and Conference Proceedings Citation Index (CPCI). We did not use the Emerging Sources Citation Index, because this citation index is not included in the WoS license of CWTS.

Scopus uses the same funder registry as Crossref, making it relatively easy to compare the funding data available in Crossref and Scopus. WoS takes a different approach and uses its own funder registry. This registry provides a unified name to each funder. Due to the different

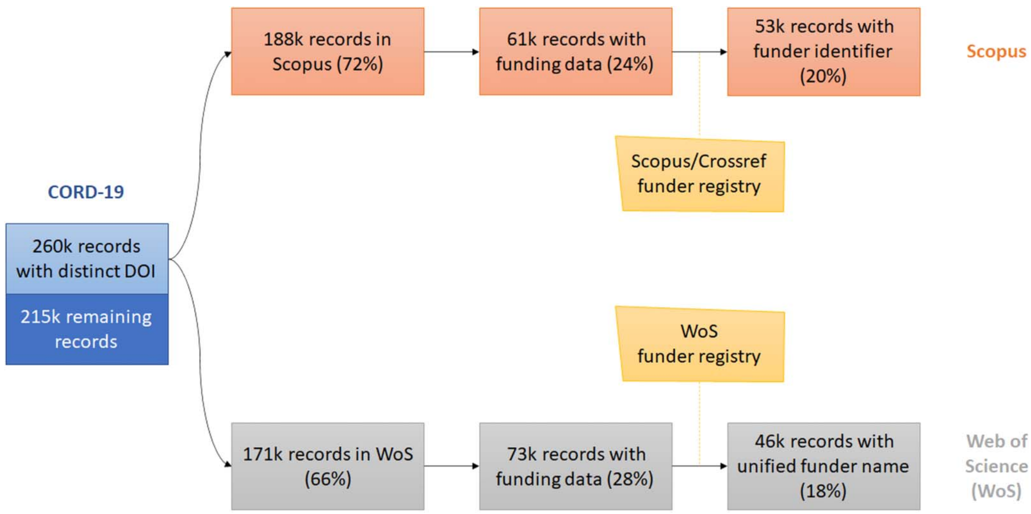

Figure 5. Funding data for COVID-19 publications in Scopus and WoS.

approach taken by WoS, there is no easy way to compare Crossref and WoS in terms of the availability of funding data at the level of individual funders. We therefore compare Crossref and WoS only in terms of whether publications do or do not report funding, without taking into account the funder that provided the funding.

In this section, we first analyze the availability of funding data in Scopus and WoS. We then look at the top funders and finally we explore the differences between Scopus, WoS, and Crossref.

4.1. Availability of Funding Data

Bibliometric databases have different scopes due to differences in their inclusion criteria. Visser, van Eck, and Waltman (2021) recently reported that overall WoS covers fewer publications than Scopus, even though there are some publications, for instance meeting abstracts and book reviews, that are covered by WoS and not by Scopus. For DOIs in CORD-19, Figure 5 shows that Scopus indexes more publications than WoS (187,518 vs. 171,130). However, looking at publications with funding data, Scopus has lower coverage than WoS (61,168 vs. 73,444 publications). On the other hand, Scopus has higher coverage than WoS (52,747 vs. 46,070 publications) if we consider only publications that include funder identifiers or unified funder names.

Both in Scopus and in WoS, the availability of funding data is higher than in Crossref, where funding data are available for 44,820 publications, of which 36,008 include funder identifiers (see Figure 1).

4.2. Top Funders

Scopus uses the Crossref Funder Registry described in Section 2.2.2, while WoS has its own registry of funders. On account of this, we do not make a direct comparison of the number of publications per funder in Scopus and WoS. Instead, we present separate statistics for each of the two databases.

For Scopus and WoS, Table 3 shows the top funders in terms of the number of COVID-19 publications. For Scopus, we look at funders at the highest level of the Funder Registry

Table 3. Top 30 funders of COVID-19 publications (based on Scopus and WoS data)

Scopus		WoS	
Funder	# pub.	Funder	# pub.
U.S. Department of Health and Human Services	13,559	United States Department of Health & Human Services	11,857
National Natural Science Foundation of China	5,938	National Institutes of Health (NIH) - USA	11,341
European Commission	3,373	National Natural Science Foundation of China (NSFC)	7,946
Ministry of Science and Technology of the People's Republic of China	2,380	European Commission	4,875
UK Research and Innovation	2,269	NIH National Institute of Allergy & Infectious Diseases (NIAID)	1,832
National Science Foundation (US)	1,587	National Science Foundation (NSF)	1,464
Ministry of Education, Culture, Sports, Science and Technology (Japan)	1,585	German Research Foundation (DFG)	1,424
Government of Canada	1,454	Canadian Institutes of Health Research (CIHR)	1,305
Deutsche Forschungsgemeinschaft	1,146	Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT)	1,284
National Institute for Health Research (UK)	1,118	Medical Research Council UK (MRC)	1,254
Wellcome Trust	1,022	UK Research & Innovation (UKRI)	1,221
Ministry of Education of the People's Republic of China	972	Wellcome Trust	1,201
Pfizer	866	National Council for Scientific and Technological Development (CNPq)	1,022
Ministério da Ciência, Tecnologia e Inovação (Brazil)	819	National Health and Medical Research Council of Australia	1,012
U.S. Department of Defense	798	Japan Society for the Promotion of Science	901
Department of Health, Australian Government (Australia)	748	CAPES (Brazil)	740
National Research Foundation of Korea	697	Fundamental Research Funds for the Central Universities	739
Ministerio de Economía y Competitividad (Spain)	658	National Institute for Health Research (NIHR)	736
Ministry of Finance (China)	633	NIH National Cancer Institute (NCI)	722
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Brazil)	609	NIH National Heart Lung & Blood Institute (NHLBI)	685
Merck	582	French National Research Agency (ANR)	668

Bill and Melinda Gates Foundation	580	Natural Sciences and Engineering Research Council of Canada (NSERC)	641
AstraZeneca	577	United States Department of Defense	622
Novartis	576	Bill & Melinda Gates Foundation	620
Roche	567	European Research Council (ERC)	578
COVIDien	551	National Basic Research Program of China	560
Auris Health	549	CGIAR	555
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung	532	NIH National Institute of General Medical Sciences (NIGMS)	552
Chinese Academy of Sciences	505	Swiss National Science Foundation (SNSF)	538
Medtronic	483	Ministry of Science and Technology, Taiwan	533

hierarchy, following the approach discussed in Section 2.2.2. For WoS, we use unified funder names.

Table 3 provides three interesting insights:

- First, the table shows that by and large, the top funders are the same in Scopus and WoS and match those in the open data obtained from Crossref. In fact, the first three funders are the same across the three databases: the U.S. Department of Health & Human Services, followed by the National Natural Science Foundation of the China (NSFC) and the European Commission. Other top funders listed in Table 1 based on Crossref are also visible among the top funders listed in Table 3 based on the proprietary databases.
- A second observation is the difficulty of making comparisons between databases that use different registries of funders. WoS harmonizes funder names, but unlike Scopus and Crossref, it does not enable funders to be aggregated into higher level entities. While one may intuitively infer that the different institutes of the NIH belong to the same higher level entity, it requires considerable knowledge of the funding landscape to know that the European Research Council, which is listed as a separate organization in the case of WoS, is part of the European Commission in the case of Scopus and Crossref. Hence, when comparing funding data from different databases, it is essential to pay close attention to the way in which relations between funding entities are handled.
- A third observation relates to pharmaceutical companies, which feature prominently on the Scopus list but not on the Crossref and WoS lists.

We explore the differences between the three databases in more detailed in the next section.

4.3. Differences Between the Databases

In the previous sections, we analyzed differences at an aggregate level in funding data obtained from three different databases: Crossref, Scopus, and WoS. In this section, we present an analysis at a more detailed level, focusing on the extent to which—at the level of individual publications—funding data obtained from these three databases differs or overlaps.

4.3.1. Intersections and differences

Figure 6 shows the overlap and the differences between the three databases in terms of publications for which funding data are available. The relatively small overlap is remarkable. There are 95,292 publications with funding data in at least one of the three databases. Only 23,950 of these publications have funding data in all three databases, an overlap of 25%. The number of publications with funding data in only one of the databases is largest for WoS (16,155). It is somewhat smaller for Scopus (12,738) and smallest for Crossref (6,209).

The differences shown in Figure 6 are partly due to differences in the publications indexed in the three databases. As indicated in Figure 5, of the DOIs in CORD-19, only 72% can be linked to publications indexed in Scopus and only 66% to publications indexed in WoS. In contrast, 98% of the DOIs can be linked to publications in Crossref, as shown in Figure 1.

In Figure 7, we therefore restrict the analysis to the 141,291 publications indexed in all three databases. Of these publications, 72,402 have funding data in at least one of the databases and 23,950 have funding data in all three databases, resulting in an overlap of 33%. As

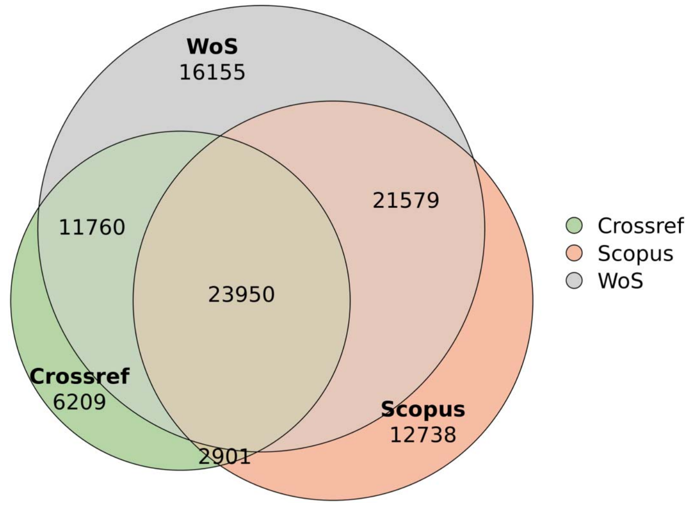

Figure 6. Overlap of Crossref, Scopus, and WoS in terms of COVID-19 publications with funding data (considering all publications indexed in at least one of the three databases).

in Figure 6, the number of publications that have funding data in one database but not in the others is largest for WoS (11,714). It is somewhat smaller for Scopus (8,457) and smallest for Crossref (729).

4.3.2. *Accuracy of funding data*

We now analyze the accuracy of funding data for individual publications by comparing funding data obtained from the different databases with funding information found in the full text of publications. The comparison is based on a stratified random sample of 120 publications. For each of the three databases considered, we drew a random sample of 40 publications that have funding data only in that database and not in the other two databases. In addition, given

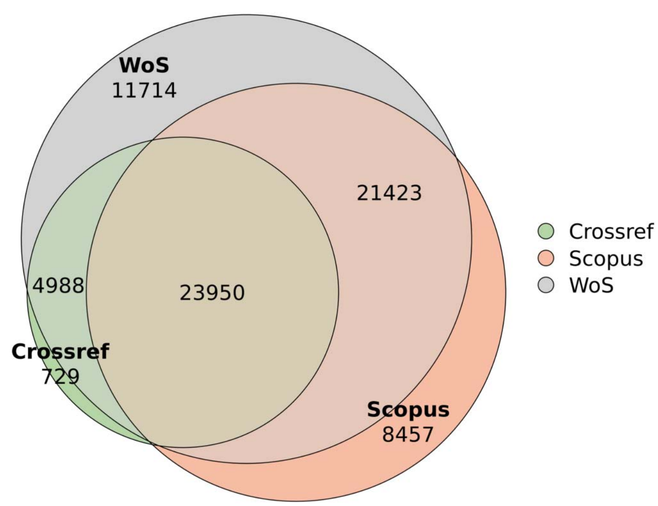

Figure 7. Overlap of Crossref, Scopus, and WoS in terms of COVID-19 publications with funding data (considering only publications indexed in all three databases).

the notable presence of pharmaceutical companies among the funders of COVID-19 publications in Scopus, we also analyze a random sample of 25 publications that have at least one private entity as funder in Scopus (i.e., an entity classified in the Funder Registry as "private" and "for-profit company/industry"). The samples are available in the Supplementary material (Mugabushaka et al., 2022).

Our findings can be summarized as follows:

- In the *Crossref sample*, we found the funding data for 27 of the 40 publications to correspond to the text in either the acknowledgment or the funding information section of a publication. In the following we refer to this as "correct" funding data. The correct entries include two instances in which the funding information is in fact a statement that there was no specific funding supporting the work (e.g., "This work received no specific grant from a funding agency"). Although the sample is too small to generalize to Crossref as a whole, this type of funding statement may be an important one that deserves further analysis. Databases that provide funding data may consider including it in their taxonomies. Of the 13 publications with incorrect funding data, four were apparently due to an error of the extraction algorithm, which for example mistook the affiliation of the authors for a funding body. In one case, the funding information was partially correct: One funding organization listed in the acknowledgment was missed but others correctly identified. In the other eight cases, the funding information could not be located anywhere in the full text of the publication.
- In the *Scopus sample*, for 15 of the 40 publications, the funding data corresponds to the funding statement found in the full text. Twenty-five cases were found to be errors, most probably of the algorithm for extracting funding information. The most common error was the algorithm incorrectly identifying the section of a publication that includes a funding statement. Sometimes a conflict-of-interest section was incorrectly interpreted as a funding statement. In other cases, the acknowledgment section was interpreted as a funding statement, when the publication in fact included a separate funding information section. Other errors, for example, included mistaking a natural person thanked by the authors for a funding body, or interpreting the affiliation of a researcher mentioned in the acknowledgment section as a funding body. In four cases, the funding information could not be found anywhere in the full text of the publication[3].
- In the *WoS sample*, we found that in 37 of the 40 cases the funding data corresponds to the funding statement in the full text. In two of these cases, we noted that the funding information provided in the full text was ambiguous. The funding information section stated that there was no funding to report, but the acknowledgment section mentioned a funding body that provided financial support. In the three cases in which no relevant funding statement could be found in the full text, there was an error, most probably caused by the extraction algorithm mistaking a conflict-of-interest section for a funding statement (see also Grassano et al., 2017; Lewison & Sullivan, 2015).
- In the *sample of publications with Scopus funding data that includes a private entity*, five of the 25 publications indeed contained a funding statement mentioning the private entity. In the other 20 cases, the funding data was incorrect: The private entity was

[3] Scopus informed us that its funding data should be seen as "work in progress" as there are still many incremental improvements in the planning. Currently Scopus focuses on optimizing the data for the top 300 funders.

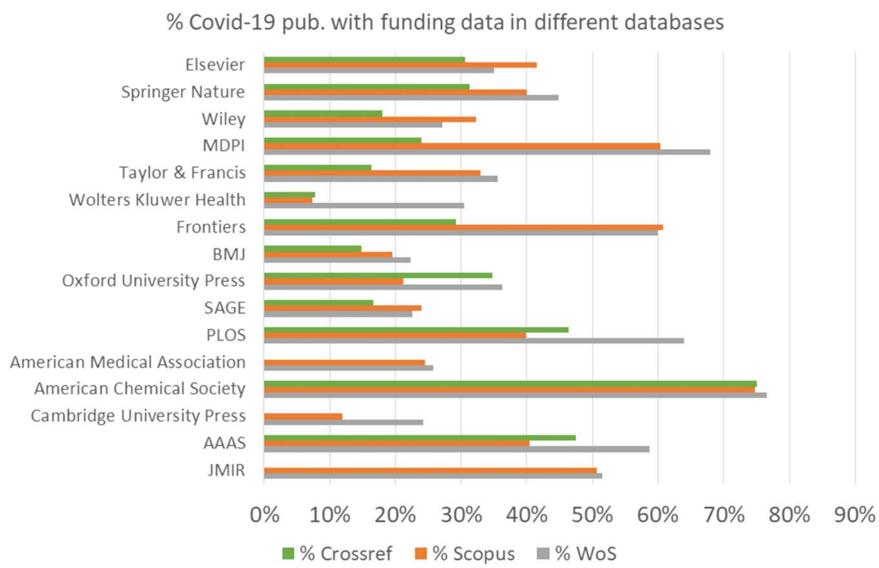

Figure 8. Percentage of COVID-19 publications with funding data, breakdown by publisher and database (considering only publications indexed in all three databases).

mentioned in the conflict-of-interest section, but not as a funder of the research. In Scopus, this problem occurred in all 20 publications. It also occurred in four publications in WoS and in one publication in Crossref.

4.3.3. Differences by publisher

We now turn to differences by publisher in the coverage of funding data. Figure 8 includes—for each of the databases considered—the share of publications with funding data. We restricted the analysis to publications from 2020 and 2021 and show only publishers with 500 or more COVID-19 publications. Statistics for other publishers are available in the Supplementary material (Mugabushaka et al., 2022).

Figure 8 shows some interesting differences between Scopus and WoS, but these are not as striking as the differences between the two proprietary databases and Crossref.

Two publishers, Oxford University Press and American Chemical Society, do an excellent job of submitting funding data to Crossref. For these publishers, the number of publications with funding data in Crossref is almost as large as in WoS and, in the case of Oxford University Press, substantially larger than in Scopus.

For many other publishers, the number of publications with funding data in Crossref is considerably below the corresponding number in WoS, and in most cases also below the corresponding number in Scopus. These publishers seem to have gaps in the funding data they submit to Crossref, or they may have started submitting funding data only recently.

Figure 8 also reveals three publishers that do not submit funding data to Crossref at all: American Medical Association, Cambridge University Press, and JMIR.

There is a need to better understand why some publishers include funding information in the metadata they provide to Crossref while others do not. A possible explanation is that awareness among publishers of the importance of submitting funding data to Crossref still needs to grow, and some publishers may also need more time to implement the submission of funding data throughout all their workflows.

5. SUMMARY AND OUTLOOK

The COVID-19 pandemic has turned the world upside down. As the surge of cases threatens to overwhelm healthcare systems and the death toll increases around the world, the effects of containment measures reverberate through economies and societies, sending shock waves that many fear will be also felt in years to come. Researchers have been among the first professions contributing to tackling the pandemic, and research funding organizations have adapted their programs or developed new ones to support them.

As a result, there has been an explosion of scientific papers on all aspects of the pandemic, by some accounts making up 4% of the 2020 scientific output indexed by major bibliometric databases. This has led also to scientometric analyses trying to uncover patterns and trends in this vast literature. In this paper we have looked into one aspect that has received little attention so far: the funding of COVID-19 research. It is important to understand how past and current funding have contributed to tackling the pandemic. Such an understanding can not only inform the design of adequate mechanisms for future emergencies but can also help funders meet their accountability obligations. As the pandemic has shown, scientific evidence that generates sound knowledge on which solutions and public policies are built has to compete with well-organized disinformation campaigns. The disclosure of funding sources can also enhance the transparency of the research process and increase the public's confidence in scientific findings.

The main objective of this paper was to explore the extent to which openly available data sets (Crossref funding data) can help in the study of funding of COVID-19 research (operationalized by the CORD-19 data set). We also aimed to make a comparison with the availability of funding data in proprietary bibliometric databases (Scopus and WoS).

We found that only 17% of the CORD-19 publications with DOIs have funding data in Crossref. This rate was higher for the proprietary databases: 24% for Scopus and 28% for WoS. Considering only publications indexed by a database, we found that 33% of the CORD-19 publications indexed in Scopus have funding data. The corresponding share for WoS is 43%.

In terms of the main funders of COVID-19 research, the three databases paint a broadly similar picture. The three funders with most publications in CORD-19 are the U.S. Department of Health and Human Services (mainly the NIH), the National Natural Science Foundation of China, and the European Commission. There are some differences in lower ranks.

By comparing publications with funding data in the three databases, we found a relatively low overlap. Considering only publications indexed in all three databases, only 33% of the publications with funding data in at least one database have funding data in all three databases. For the two proprietary databases, the overlap is 64%.

We also assessed, based on small samples, the accuracy of the funding data present in one database but not in the other two. Our analysis shows that most funding data exclusive to WoS (i.e., data available neither in Crossref nor in Scopus) matched with funding information in the full text of publications. The share of publications for which we could not confirm the correctness of the funding data based on the full text of the publication was also relatively low in Crossref, but it was quite high in Scopus.

After observing that the list of top funders from Scopus includes more pharmaceutical companies than the lists from Crossref and WoS, we also checked manually a sample of 25 publications which had a company among the funding organizations. We found that in most cases this did not correspond to the funding information included in the paper. Rather, it seems to be the result of an error made by the algorithm used to extract and structure funding data. In most cases, the algorithm incorrectly treats the conflict of interest or disclosure section of a paper as a funding statement. As a result, pharmaceutical companies are often presented as funders of the research presented in a paper, while in fact they are funders or collaborators in other activities of the authors[4].

The main observation from this study is the *limited coverage of funding data* in open data infrastructures. Using Crossref funding data alone allows us to paint only a partial picture of who is funding COVID-19 research. In comparison with proprietary databases, the share of CORD-19 publications (with DOI) that have funding data in Crossref is about seven percentage points lower than in Scopus and about 11 percentage points lower than in WoS. The limited coverage of funding data in Crossref can be explained by differences in the metadata deposited by publishers to Crossref. Although for some publishers we have nearly full coverage of funding data, for others the coverage is relatively low, and there are also publishers that do not deposit funding data at all. We also observed that in proprietary databases the coverage by publisher rarely exceeds 75%. Publications without funding data may present research for which the authors did not receive any funding. However, it is also possible that the authors did receive funding and did report this in their publication, but that the funding information was not processed properly due to algorithmic mistakes.

A second observation from this study is the *uncertain quality of funding data*. As already mentioned, in a small random sample that we analyzed, we found that most funding data related to pharmaceutical companies is based on algorithmic errors, mainly because extraction algorithms confuse conflict of interest statements with funding statements. This issue affects all databases considered but seems more severe in Scopus.

In the following we offer some reflections on how the availability of funding data can be improved.

Authors provide funding information in their papers to comply with the requirements of publishers and funders. On the one hand, publishers' ethical guidelines increasingly require disclosure of funding sources. In the case of research with commercial or political interests at stake, this transparency helps readers assess the extent to which the credibility of the findings may be related to possible conflicts of interests of the authors. This is standard practice in many medical journals. On the other hand, almost all major research funders require—often as part of their grant terms and conditions—that grant holders explicitly acknowledge the funder's support in publications to which the funding has contributed.

By collecting and making available funding data provided in funding statements in publications, publishers provide an important service to various stakeholders, including the following:

- to readers by providing transparency on the sponsor of a study;
- to authors by helping them comply with funders' requirements;
- to funders by allowing them to easily identify the results of their funding; and
- to scientometricians by enabling them to study the effectiveness and impact of funding practices.

However, as the results presented in this paper show, funding data in open infrastructures is lacking in terms of coverage and to some extent also quality.

[4] Scopus informed us that it currently focuses on optimizing its funding data for the top 300 funders. While this may be good enough for global comparisons, our findings show that it may give an inaccurate picture in more specific analyses.

Publishers should be encouraged to sustain and intensify their efforts to submit funding data to Crossref. One improvement that could be considered is to extend the data that can be provided to Crossref to also include the raw funding information text. This would allow publishers who cannot commit resources to extract structured funding data from papers to participate in this effort, overcoming the current situation in which publishers need to choose between providing structured funding data to Crossref and providing no funding data at all. Another advantage would also be that having the funding statements as provided in papers could help improve the quality of funding data. As better algorithms become available, they can be applied, also retrospectively, to the available funding statements to turn these statements into high-quality structured funding data.

Funders also have an important role in improving both the availability and the quality of funding data in open data infrastructures. In particular, funders should support the efforts of open scholarly infrastructures to create persistent identifiers. The Funder Registry used by Crossref and Scopus and the deposition of funding data to Crossref should be seen as part of broader efforts to improve the availability of high-quality open funding data, building on past efforts—such as the guidelines of the UK Research Information Network (RIN, 2008), and continuing to evolve. Crossref recently started an initiative to assign DOIs to research grants[5]. Funders should take up this opportunity and offer guidance to grant holders on how persistent identifiers for grants should be used in funding statements in publications.

As scientific results increasingly impact our daily life, researchers, funders, publishers, research organizations, and society at large share an interest in safeguarding the trust and credibility that scholar communication enjoys. Working together on realizing high-quality open funding data is an essential step in this endeavor.

ACKNOWLEDGMENTS

We are grateful to Clara Calero, Dan Gibson, and Jeroen van Honk (CWTS, Leiden University), Ginny Hendricks (Crossref), M'hamed El Aisati (Elsevier Scopus), and Gali Halevi (Clarivate WoS) for their feedback on an earlier draft of this paper. We also thank two anonymous reviewers for their helpful comments and suggestions.

AUTHOR CONTRIBUTIONS

Alexis-Michel Mugabushaka: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Visualization; Writing—original draft. Nees Jan van Eck: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Visualization; Writing—review & editing. Ludo Waltman: Conceptualization; Visualization; Writing—review & editing.

COMPETING INTERESTS

From 2018 until 2020, Alexis-Michel Mugabushaka was a member of Crossref's Funders Advisory Group on persistent identifiers for grants. In addition, the work was performed while he was on secondment to the Directorate-General for Research and Innovation (DG RTD) of the European Commission (EC). The views expressed in this paper are the authors'. They do not reflect official positions of the EC or the European Research Council.

FUNDING INFORMATION

The authors did not receive any funding for the research reported in this paper.

[5] https://www.crossref.org/blog/request-for-feedback-on-grant-identifier-metadata/.

DATA AVAILABILITY

The version of the CORD-19 data set used in this paper was released by Microsoft Academic on February 22, 2021 and corresponds to the release from February 15, 2021 by the Allen Institute. It is accessible at https://magcord19.blob.core.windows.net/mapping/2021-02-22-CORD-19-MappedTo-2021-02-15-MAG-Backfill.csv.

The mapping of funding organizations to the corresponding top-level entities (see Section 2.2.2) is available in Zenodo (van Eck & Mugabushaka, 2021): https://doi.org/10.5281/zenodo.5562841.

The Supplementary material is also available in Zenodo (Mugabushaka et al., 2022): https://doi.org/10.5281/zenodo.6805409.

REFERENCES

Álvarez-Bornstein, B., & Montesi, M. (2020). Funding acknowledgements in scientific publications: A literature review. *Research Evaluation*, *29*(4), 469–488. https://doi.org/10.1093/reseval/rvaa038

Álvarez-Bornstein, B., Morillo, F., & Bordons, M. (2017). Funding acknowledgments in the Web of Science: Completeness and accuracy of collected data. *Scientometrics*, *112*(3), 1793–1812. https://doi.org/10.1007/s11192-017-2453-4

Andersen, J. P., Nielsen, M. W., Simone, N. L., Lewiss, R. E., & Jagsi, R. (2020). COVID-19 medical papers have fewer women first authors than expected. *eLife*, *9*, e58807. https://doi.org/10.7554/eLife.58807, PubMed: 32538780

Aristovnik, A., Ravšelj, D., & Umek, L. (2020). A bibliometric analysis of COVID-19 across science and social science research landscape. *Sustainability*, *12*(21), 9132. https://doi.org/10.3390/su12219132

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, *1*(1), 377–386. https://doi.org/10.1162/qss_a_00019

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, *1*(1), 363–376. https://doi.org/10.1162/qss_a_00018

Bleck, T. P., Buchman, T. G., Dellinger, R. P., Deutschman, C. S., Marshall, J. C., ... Zimmerman, J. J. (2020). Pandemic-related submissions: The challenge of discerning signal amidst noise. *Critical Care Medicine*, *48*(8), 1099–1102. https://doi.org/10.1097/CCM.0000000000004477, PubMed: 32697478

Brainard, J. (2020). New tools aim to tame pandemic paper tsunami. *Science*, *368*, 924–925. https://doi.org/10.1126/science.368.6494.924, PubMed: 32467369

Colavizza, G., Costas, R., Traag, V. A., van Eck, N. J., van Leeuwen, T., & Waltman, L. (2021). A scientometric overview of CORD-19. *PLOS ONE*, *16*(1), e0244839. https://doi.org/10.1371/journal.pone.0244839, PubMed: 33411846

Cross, S., Rho, Y., Reddy, H., Pepperrell, T., Rodgers, F., ... Keestra, S. (2021). Who funded the research behind the Oxford-AstraZeneca COVID-19 vaccine? *BMJ Global Health*, *6*, e007321. https://doi.org/10.1136/bmjgh-2021-007321, PubMed: 34937701

Else, H. (2020). How a torrent of COVID science changed research publishing—In seven charts. *Nature*, *588*, 553. https://doi.org/10.1038/d41586-020-03564-y, PubMed: 33328621

Grassano, N., Rotolo, D., Hutton, J., Lang, F., & Hopkins, M. M. (2017). Funding data from publication acknowledgments: Coverage, uses, and limitations. *Journal of the Association for Information Science and Technology*, *68*(4), 999–1017. https://doi.org/10.1002/asi.23737

Hao, K. (2020). The scientists and technologists who dropped everything to fight covid-19. *MIT Technology Review*, April 15. https://www.technologyreview.com/2020/04/15/999478/scientists-engineers-volunteer-fight-covid-19-pandemic/

Hendricks, G., Tkaczyk, D., Lin, J., & Feeney, P. (2020). Crossref: The sustainable source of community-owned scholarly metadata. *Quantitative Science Studies*, *1*(1), 414–427. https://doi.org/10.1162/qss_a_00022

Herzog, C., Hook, D., & Konkiel, S. (2020). Dimensions: Bringing down barriers between scientometricians and data. *Quantitative Science Studies*, *1*(1), 387–395. https://doi.org/10.1162/qss_a_00020

Hossain, M., Sarwar, S., McKyer, E. L., & Ma, P. (2020). Applications of artificial intelligence technologies in COVID-19 research: A bibliometric study. *Preprints*, *2020*, 2020060161. https://doi.org/10.20944/preprints202006.0161.v1

Hurst, P., & Greaves, S. (2021). COVID-19 Rapid Review cross-publisher initiative: What we have learned and what we are going to do next. *Learned Publishing*, *34*(3), 450–453. https://doi.org/10.1002/leap.1375, PubMed: 34230774

Ioannidis, J. P. A., Salholz-Hillel, M., Boyack, K. W., & Baas, J. (2021). The rapid, massive growth of COVID-19 authors in the scientific literature. *Royal Society Open Science*, *8*(9), 210389. https://doi.org/10.1098/rsos.210389, PubMed: 34527271

Kanakia, A., Wang, K., Dong, Y., Xie, B., Lo, K., ... Wu, C. H. (2020). Mitigating biases in CORD-19 for analyzing COVID-19 literature. *Frontiers in Research Metrics and Analytics*, *5*, 596624. https://doi.org/10.3389/frma.2020.596624, PubMed: 33870059

Kiszewski, A. E., Cleary, E. G., Jackson, M. J., & Ledley, F. D. (2021). NIH funding for vaccine readiness before the COVID-19 pandemic. *Vaccine*, *39*(17), 2458–2466. https://doi.org/10.1016/j.vaccine.2021.03.022, PubMed: 33781600

Kokol, P., & Blažun Vošner, H. (2018). Discrepancies among Scopus, Web of Science, and PubMed coverage of funding information in medical journal articles. *Journal of the Medical Library Association*, *106*(1), 81–86. https://doi.org/10.5195/jmla.2018.181, PubMed: 29339937

Kwon, D. (2020). Scientists around the globe pivot their research to SARS-CoV-2. *The Scientist*, April 6. https://www.the-scientist.com/news-opinion/scientists-around-the-globe-pivot-their-research-to-sars-cov-2-67385

Lewison, G., & Sullivan, R. (2015). Conflicts of interest statements on biomedical papers. *Scientometrics*, *102*(3), 2151–2159. https://doi.org/10.1007/s11192-014-1507-0
Liu, W. (2020). Accuracy of funding information in Scopus: A comparative case study. *Scientometrics*, *124*(1), 803–811. https://doi.org/10.1007/s11192-020-03458-w
Liu, W., Tang, L., & Hu, G. (2020). Funding information in Web of Science: An updated overview. *Scientometrics*, *122*(3), 1509–1524. https://doi.org/10.1007/s11192-020-03362-3
Mohadab, M. E., Bouikhalene, B., & Safi, S. (2020). Bibliometric method for mapping the state of the art of scientific production in Covid-19. *Chaos, Solitons, and Fractals*, *139*, 110052. https://doi.org/10.1016/j.chaos.2020.110052, PubMed: 32834606
Mugabushaka, A.-M. (2020). Linking publications to funding at project level: A curated dataset of publications reported by FP7 projects. *arXiv*, arXiv:2011.07880. https://doi.org/10.48550/arXiv.2011.07880
Mugabushaka, A.-M., van Eck, N. J., & Waltman, L. (2022). Funding Covid-19 research: Insights from an exploratory analysis using open data infrastructures—Supplementary material [Data set]. *Zenodo*. https://doi.org/10.5281/zenodo.6112762
Paul-Hus, A., Desrochers, N., & Costas, R. (2016). Characterization, description, and considerations for the use of funding acknowledgement data in Web of Science. *Scientometrics*, *108*(1), 167–182. https://doi.org/10.1007/s11192-016-1953-y
RIN. (2008). *Acknowledgement of funders in scholarly journal articles: Guidance for UK research funders, authors and publishers*. Retrieved from https://www.ukri.org/wp-content/uploads/2020/10/RIN-251020-FundersAcknowledgementInScholarlyjournalArticles.pdf
Shorten, C., Khoshgoftaar, T. M., & Furht, B. (2021). Deep Learning applications for COVID-19. *Journal of Big Data*, *8*(1), 18. https://doi.org/10.1186/s40537-020-00392-9, PubMed: 33457181
Shueb, S., Gul, S., Nisa, N. T., Shabir, T., Rehman, S. U., & Hussain, A. (2022). Measuring the funding landscape of COVID-19 research. *Library Hi Tech*, *40*(2), 421–436. https://doi.org/10.1108/LHT-04-2021-0136
Stoye, E. (2020). How research funders are tackling coronavirus disruption. *Nature*, April 17. https://doi.org/10.1038/d41586-020-01120-2
Tang, L., Hu, G., & Liu, W. (2017). Funding acknowledgment analysis: Queries and caveats. *Journal of the Association for Information Science and Technology*, *68*(3), 790–794. https://doi.org/10.1002/asi.23713
Tao, Z., Zhou, S., Yao, R., Wen, K., Da, W., … Tao, L. (2020). COVID-19 will stimulate a new coronavirus research breakthrough: A 20-year bibliometric analysis. *Annals of Translational Medicine*, *8*(8), 528. https://doi.org/10.21037/atm.2020.04.26, PubMed: 32411751
van Eck, N. J., & Mugabushaka, A.-M. (2021). Crossref Funder Registry—Mapping to top-level funding organisations [Data set]. *Zenodo*. https://doi.org/10.5281/zenodo.5562842
van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3, PubMed: 20585380
van Eck, N. J., & Waltman, L. (2021). Crossref as a source of open bibliographic metadata. In *Proceedings of the 18th International Conference of the International Society for Scientometrics and Informetrics* (pp. 1169–1174).
Viglione, G. A. (2020). Tens of thousands of scientists are redeploying to fight coronavirus. *Nature*, March 27. https://doi.org/10.1038/d41586-020-00905-9, PubMed: 32221508
Visser, M., van Eck, N. J., & Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies*, *2*(1), 20–41. https://doi.org/10.1162/qss_a_00112
Wang, L. L., Lo, K., Chandrasekhar, Y., Reas, R., Yang, J., … Kohlmeier, S. (2020). CORD-19: The COVID-19 Open Research Dataset. *arXiv*, arXiv:2004.10706. https://doi.org/10.48550/arXiv.2004.10706